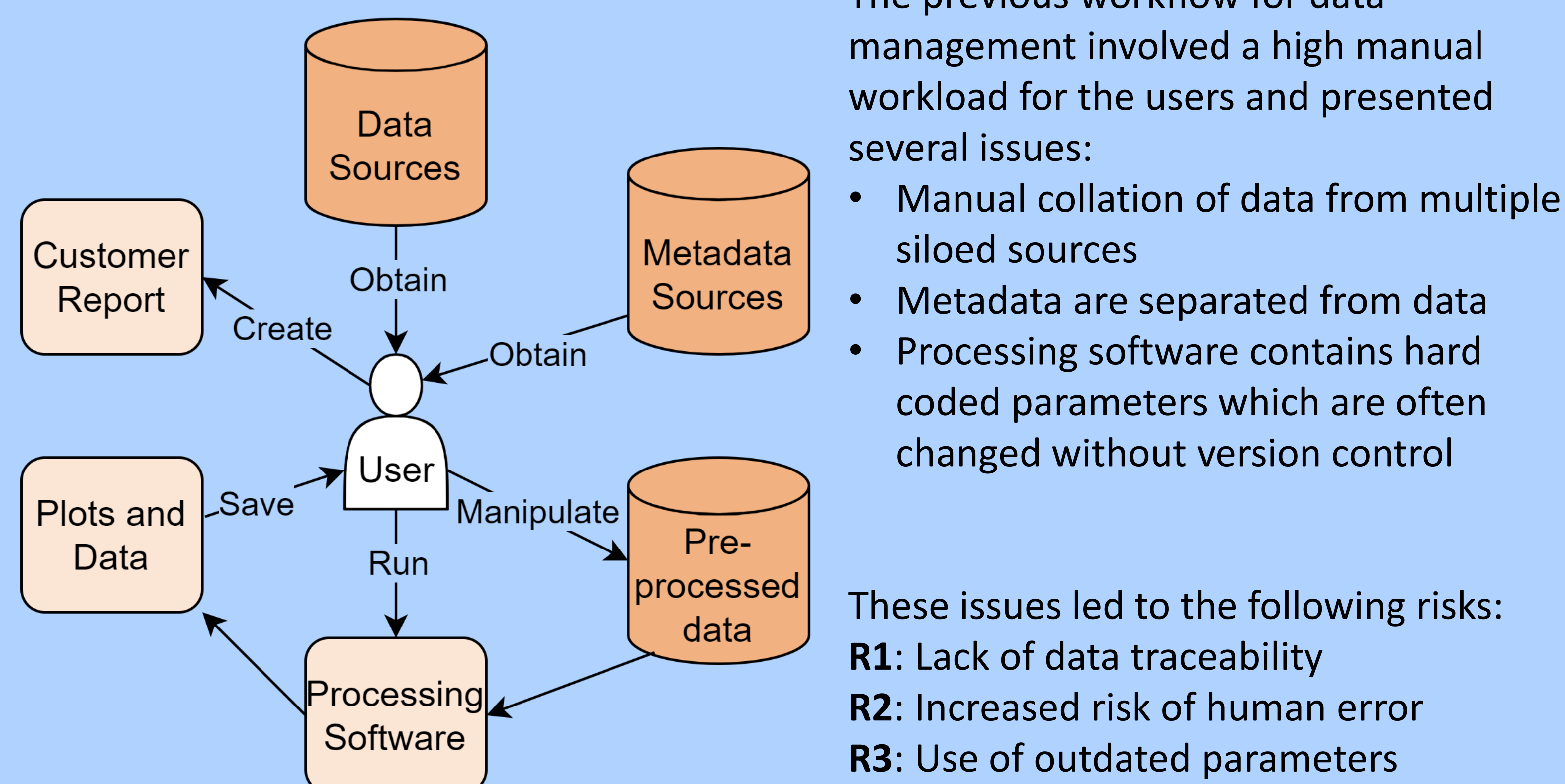


Motivation

The National Physical Laboratory is establishing a new frequency measurement service to support UK and international customers. High-quality digital data capture and provision will ensure data traceability and system resilience, while also informing future research and development of further services. We present a design a proof-of concept digitalisation system to improve and automate the existing data collection, processing and reporting.

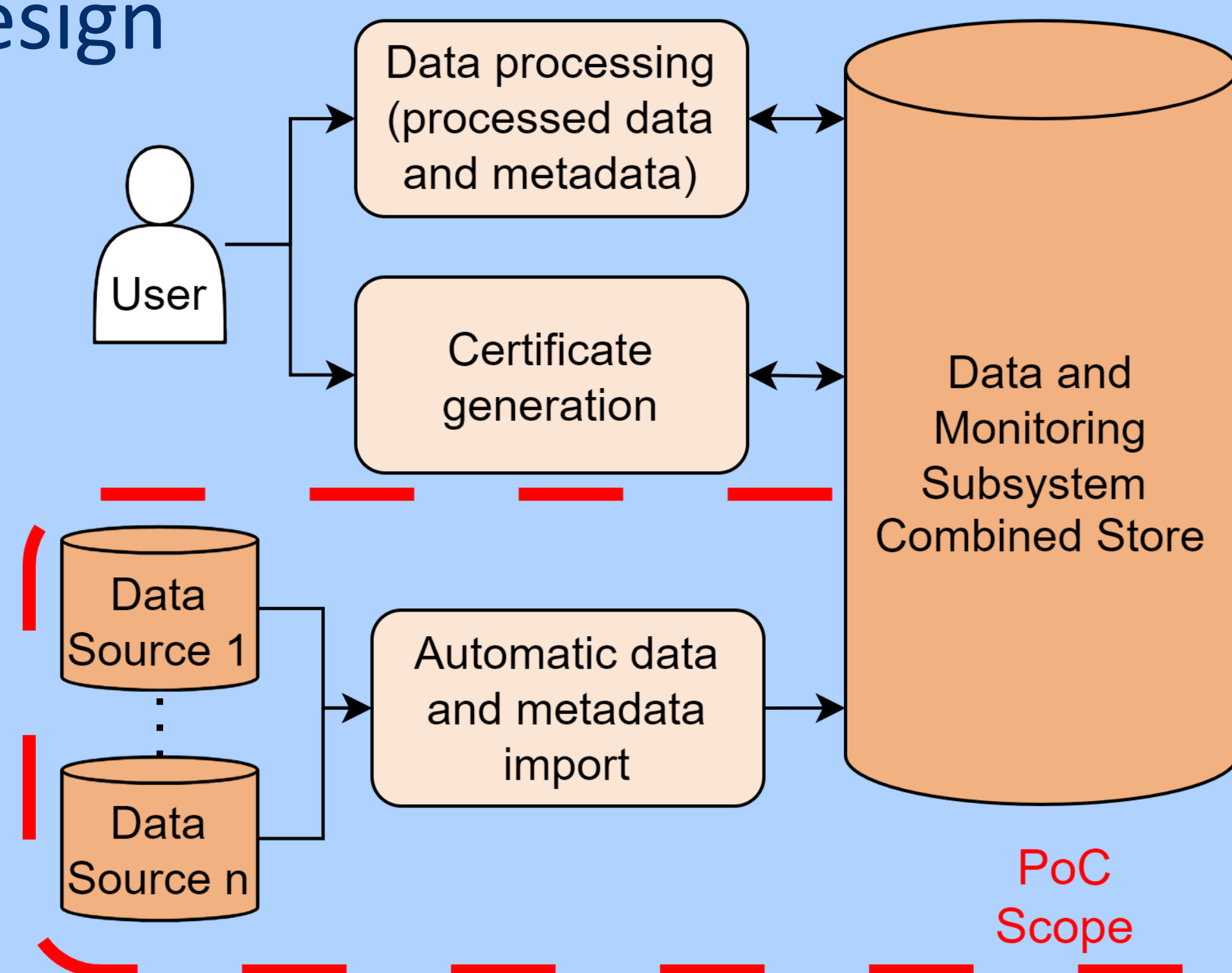
The Problem



Solution Concept Design

The concept design encapsulated four main features of the new system:

- Automatic data imports
 - Storage of data along with its metadata
 - A user interface for input of metadata
 - Automatic generation of measurement reports
- A systems engineering lifecycle was used to guarantee traceability of the solution to user requirements.



Proof of Concept Design

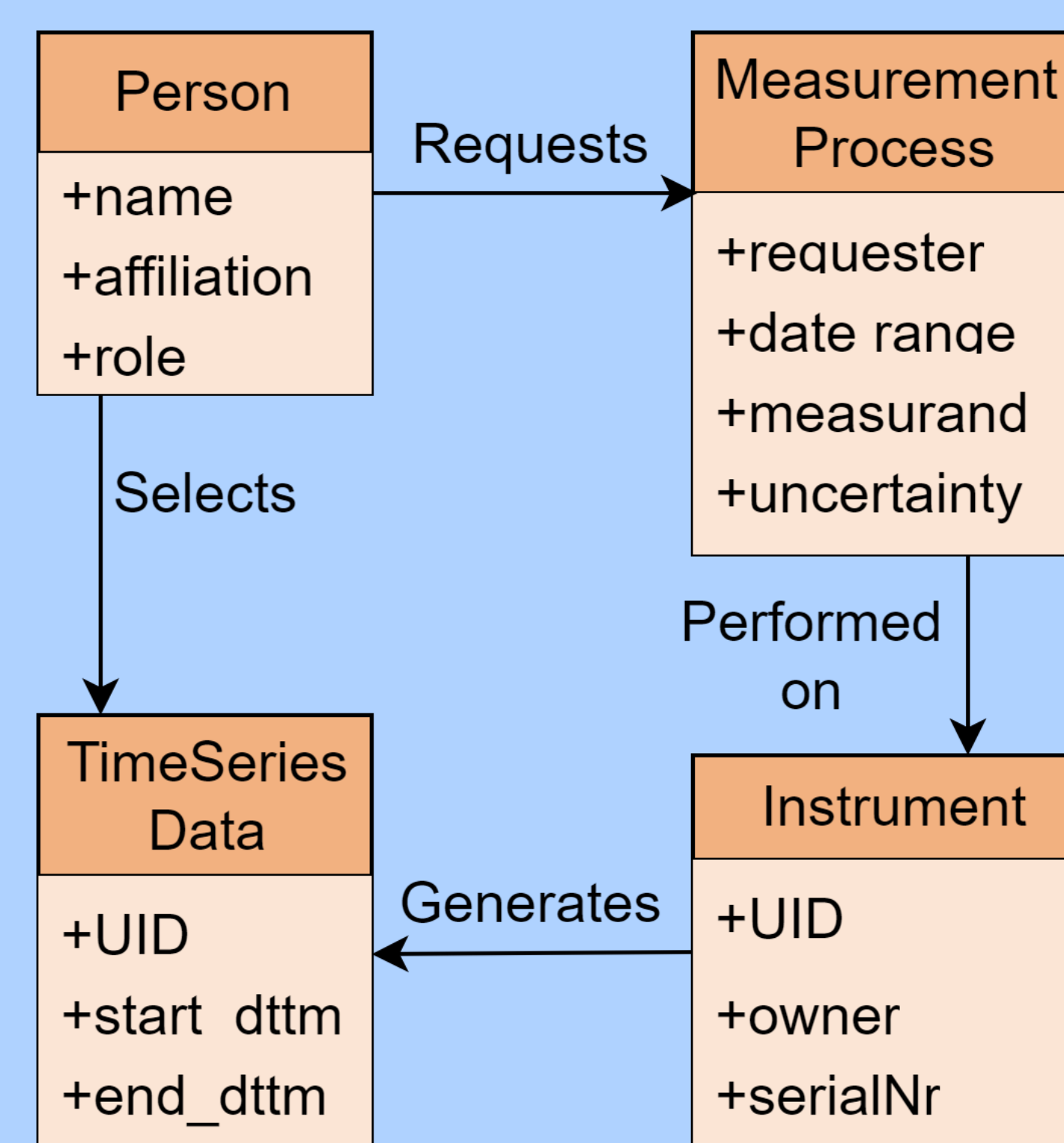
A proof of concept (PoC) data and monitoring (DM) subsystem was created to set the basis for future service digitalisation. The PoC scope covered the automated data management (capture, processing, storage and retrieval) for the radiofrequency characterisation service.

Requirements

- Scalability
- Resilience to errors
- Data persistence
- Data traceability
- FAIRness**

FAIR: compliant with Findable, Accessible, Interoperable and Reusable principles of data management.

Semantic Modelling



A high-level view of the DM subsystem metadata schema comprising an Instrument, Measurement Process, Person and Data entities is shown above. These entities were used to inform the data models used in the PoC components.

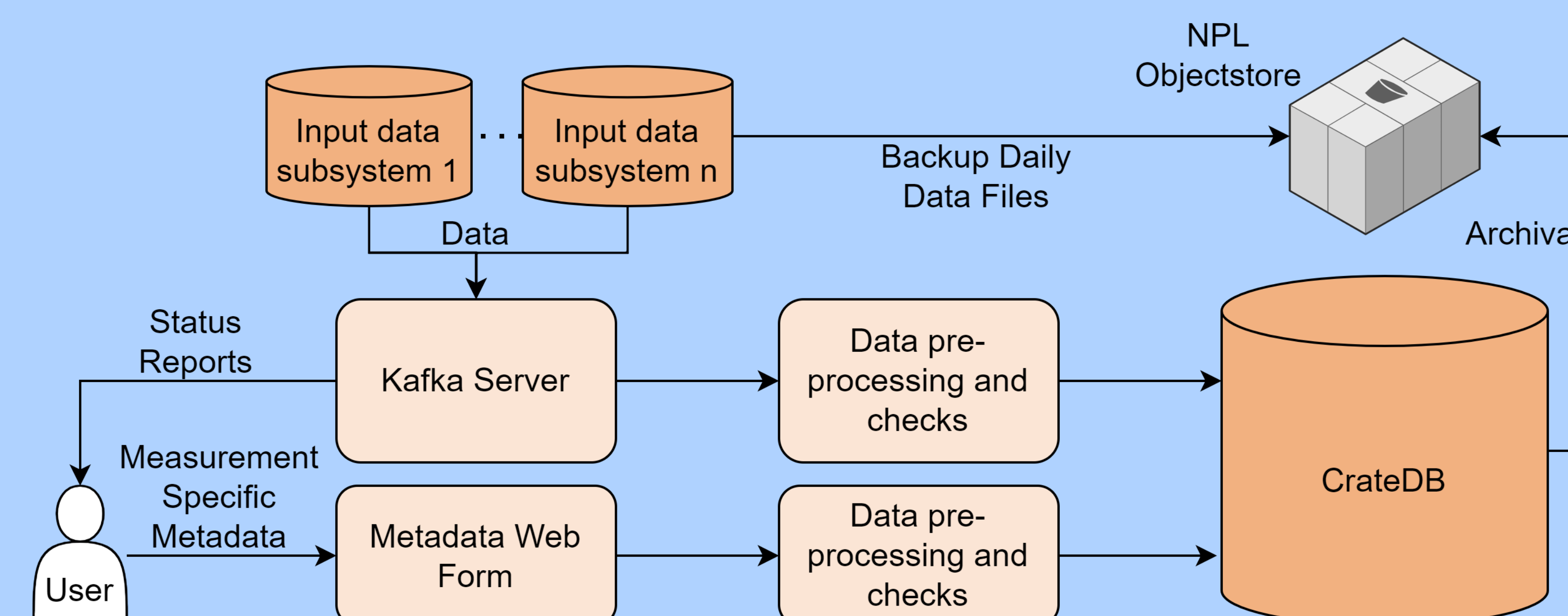
Implementation

Components

The requirements are reflected in the design choices:

- Data transfer and processing are handled by an Apache Kafka broker (covers persistence, error resilience and scalability)
- Data are stored in CrateDB, a hybrid database for data and metadata (covers FAIRness & traceability)
- Data are backed up to an S3 object store (covers persistence and scalability)

Design Solution



The PoC design solution is based on:

- Near-real-time data access via a data stream
- Measurement metadata entry via a webform
- System status monitoring through data validity and reception checks
- Data retrieval through a web user interface

Conclusions

Outcomes

This PoC design addresses the risks associated with the previous system implementation:

- Data and metadata are now stored together (**R1, R3**)
- Data can be traced through the system using unique identifiers (**R1, R2**)
- The data flow and collection are automated, removing the risk of operator error (**R2, R3**)

Future Work

- Increase the PoC technology readiness level
- Refine the semantic data models to capture metadata and align them with SI Digital Reference Point
- Add data processing capabilities
- Implement environmental monitoring
- Automate report generation